

F***ML Annotation Guidelines

Alex Reese, Kirsten Stallings, Danielle Wilson, BingYang Ye
Spring 2018

Introduction

The goal of the F***ML annotation project is to annotate the many usages and meanings of various swear words. We will be looking at 10 swear words¹ which (1) are frequently used (2) have multiple and interesting uses/meanings and (3) are swears that we believe to be *relatively* non-offensive. Choosing these 10 words creates a well-defined scope for our project, keeping the size of work manageable for our annotators and allowing us to focus and richly understand the nuances of each of these words. The choice is also an acknowledgment of the sensitive nature of our project. We are not interested in highly offensive language which targets a particular race, gender, sexual identity, or other group of people.

We hope that this annotation project would allow for the creation of a “smart” language censor. Of course a computer can easily sift through data to find and delete swear words, but with human annotation, we can capture *why* the swear was used, allowing a “smart” censor to recognize usage/meaning and replace it with an appropriate equivalent. In general, we also believe that our project will contribute greatly to the understanding of swear word usage and meanings.

Pre-processing

Each file contains 100 Reddit comments. All swear words have been pre-tagged with the OTHER tag. It is not necessary to remove the OTHER tag if the word is a swear word--just add additional tags on top of it. If a word is falsely tagged, remove the OTHER tag.

Examples of falsely tagged swear words:

My tribute website to Salt and Peppa is www.pushit.com!

*She walked down the runway with so much **sass**.*

Swear words being used in non-swear ways **do not** count as falsely tagged words.

Non-examples of falsely tagged words:

***Dick** Cheney shot a man!*

¹ See this google doc for a [list of swears](#) we are using

The *ass* chased the chickens around his pen.

In some instances, the pre-processing has tagged more than the given word. In these cases, you should still put the swear tag on only the swear word. In the example below, the preprocessing is in blue and the swear tag applied is in red:

That sucks! /n/n/r/rFuck that guy! -----> *That sucks! /n/n/r/rFuck that guy!*

Annotation Schema Overview

There are five main tags which consume only the swear word itself. No other words, whitespace, or punctuation should be a part of these five tags, even if they are compound words. For example, in *bullshit*, only *shit* is tagged.

All tags will be explained in further detail, but the following set of examples gives an overview of the schema :

1. Emphasis *...but it was pretty damn special.*
2. Reactionary *Damn it!*
3. Meaningful *The snowstorm really fucked up my plans for today.*
→ meaning subtag = “messed up”
4. Meta *I got in trouble today for saying ‘damn’.*
5. Non-Swear *The ass broke out of his pen and terrorized the chickens*

As you can see above, words tagged with MEANINGFUL will also get tagged with their exact meanings. We have therefore compiled a list of all possible meanings specific to each word. To organize and weed out meanings not associated with the given swear, these tags are broken up by swear (e.g. MEANINGFUL_DAMN). Most common meanings should be covered, but the ‘other’ tag can be chosen in cases where the annotator feels that another meaning is denoted.

Words with the EMPHASIS tag include a link tag (EMPHASIZE) to the word or phrase being emphasized, (EMPHASIZED). For example, in *That’s a big-ass pretzel*, *ass*=EMPHASIS, *big*=EMPHASIZED, and the EMPHASIZE(*ass*, *big*).

In addition to these 5 main tags, we also have the tag PHRASE to mark swear words used as part of a larger expression. This includes, for example, idiomatic expressions like “*Shit hit the fan*” and phrasal verbs like “*to fuck up*”. This tag is different from the other 5 in that it occurs in addition to the other tags. This means that the swear will still get tagged as one of the 5 main tags (consuming only the individual word), and then an additional phrase tag will

consume the entire phrase (no whitespace or punctuation on left and right boundaries) with a CONTAINS link between the two.

Emphasis

The EMPHASIS tag is given to a swear when the swear serves to emphasize another word or phrase. The test for this tag is whether the sentence could be restated without the swear word.

- (1) *There were a **shit** ton of people at the performance.*
Cf. There were a ton of people at the performance.
- (2) *What the **fuck** is going on?*
Cf. What is going on?

For every EMPHASIS tag, there must also be an EMPHASIZE tag and an EMPHASIZED tag, where EMPHASIZED is the word or phrase receiving the emphasis, and EMPHASIZE is the non-consuming link tag between the two. The examples above are tagged as follows:

- (3) *There were a **shit**(=EMPHASIS) ton(=EMPHASIZED) of people at the performance.*
- (4) *What(=EMPHASIZED) the **fuck**(=EMPHASIS) is going on?*
(=PHRASE)

(The PHRASE tag is discussed in more detail further on.)

Reactionary

The REACTIONARY tag is assigned to swears when they are used in an expressive or exclamatory way. Most often, these stand alone as their own phrase, rather than being holding meaning in a sentence. Once tagged for REACTIONARY, these swear words will also be tagged as either DIRECTED or UNDIRECTED. The following are examples of reactionary swears, broken up into directed and undirected.

Undirected:

Fuck! Daaaaamn
Shit! Holy shit!
Damn it!

Directed:

Fuck you!
Piss off!
Fuck off!

Meaningful

When a swear word is used in a way which is synonymous with a non-swear word, it receives the MEANINGFUL tag. The list of meaningful tags are specific to each word. In order to keep the meanings word-specific and organized, the word will first be tagged with the MEANINGFUL tag corresponding to the swear (e.g. MEANINGFUL_DAMN) and then tagged with one of the meanings in the drop-down list for that swear. Most common meanings should be covered, therefore the provided subtags should be used as much as possible. However, the 'other' tag can be chosen in cases where the correct meaning truly does not fit into any of the meaning subtags provided.

The following table illustrates the possible meanings for each swear along with examples:

ass	mean person	<i>"you're an ass", "you're an asshole"</i>
	fool	<i>"he made an ass of himself"</i>
	butt	<i>"my ass hurts after sitting in the car"</i>
	sex	<i>"he was looking for ass last night"</i>
bitch	mean person	<i>"you're a bitch"</i>
	feminine man	
	complain	<i>"bitch about our homework"</i>
	cool	<i>"the party was bitchin"</i>
	a woman	
	an unpleasant difficulty	<i>"this problem is a real bitch"</i>
damn	wish harm upon	
	something of little value	<i>give a damn</i>
dick	mean person	<i>"you're a dick"</i>
	penis	<i>"he got hit in the dick with a golfclub"</i>
	messing around	<i>"dicking around"</i>
douche	jerk	<i>"he's a douche"</i>

fuck	mess with	<i>"fuck with so./sth"</i>
	make a situation worse	<i>"fuck so. over"</i>
	have sex	<i>"i'd fuck him"</i>
	person	<i>"nerdy looking fuck"</i>
	messing around	<i>"fucking around/off"</i>
	care	<i>"give a fuck"</i>
	disturbing	<i>fucked up</i>
piss	pee	<i>"i need to take a piss"</i>
	angry	<i>"pissed off"</i>
pussy	wimp	<i>"they were pussies and didn't go swimming"</i>
	vagina	
shit	poop	<i>"i stepped in shit"</i>
	things/stuff	<i>"I have too much shit"</i>
	bad/poor	<i>"you look like shit"</i>
	lies/ to lie	<i>"i bullshitted that essay"</i>
	something unwanted	<i>"this is bullshit"</i>
	idiot	<i>dipshit/shithead</i>
hell	the underworld	<i>"gave me hell"</i>

Meta

The META tag is used when a swear word is used to reference the swear itself.

Examples of META swears:

Someone wrote "**BITCH**" across her windshield after she parked in two spots

My mom gets angry when I say "**fuck**" in front of my grandmother.

Non-Swear

The NON-SWEAR tag is used when a swear word is used in a way that is not a swear.

Examples of NON-SWEAR swears:

The **ass** chased the chickens around the farm.

The **bitch** had two litters of puppies

Douches are located on aisle 3 with the other hygiene products.

Dick and Jane played in the rain.

I bought my tennis racket from *Dick's*.

The sinners will be *damned* to *hell*.

Phrase

The PHRASE tag is used in addition to an individual swear tag when a swear word is used in a familiar context. Examples include phrases such as *give a damn*, *give a rat's ass*, *fucked up*, *shit hit the fan*, *dumb ass*, and *bullshit*. A good indicator of whether or not a phrase should be tagged is--if a series of words passes the tests for emphasis, reactionary, meta, meaningful, etc, the entire series of words should be tagged as a phrase in addition to the individual swear word.

Examples of phrases:

*What **the fuck** is wrong with you?*

This instance passes the test for EMPHASIS in that *the fuck* can be removed without changing the meaning but it would weaken the statement. Therefore *the fuck* should be tagged as PHRASE. NOTE: *the fuck* should not be tagged as EMPHASIS, only *fuck* should be tagged.

*My mom gets angry when I say **fuck me sideways***

In this instance, *fuck me sideways* passes the test for META in that it is referencing the words themselves. Therefore *fuck me sideways* should be tagged as PHRASE. NOTE: *fuck me sideways* should not be tagged as META, only *fuck* should be tagged.

*This is **bullshit**.*

This instance passes the test for MEANINGFUL because *bullshit* has additional meaning besides *shit*. Therefore *bullshit* should be tagged as PHRASE.

Non-examples of phrases:

*That's a **big-ass** pretzel*

In this instance, *ass* passes the test for EMPHASIS (You can take out *ass* and have a sentence with the same meaning) but *big-ass* does not (if you take out *big-ass*, you no longer know the size of the pretzel). Therefore, it is not a PHRASE.

Contains link

For each PHRASE tag, include a CONTAINS link between the phrase and the swear. For example. In *My mom gets angry when I say **fuck me sideways***, there should be a CONTAINS link between the META *fuck* and PHRASE *fuck me sideways*. In *What **the fuck** is wrong with you?*, there should be a CONTAINS link between EMPHASIS *fuck* and PHRASE *the fuck*. Since *fuck* is EMPHASIS, there will also be a EMPHASIZE link between *fuck* and *what*.